Science of the Total Environment 616–617 (2018) 678–687

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

A catchment-scale model to predict spatial and temporal burden of *E. coli* on pasture from grazing livestock

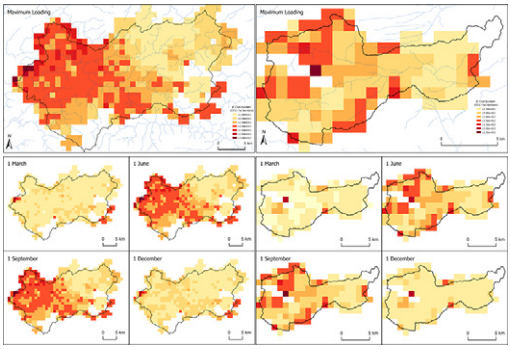
David M. Oliver [a,*], Phil J. Bartie [a], A. Louise Heathwaite [b], Sim M. Reaney [c], Jared A.Q. Parnell [a], &, Richard S. Quilliam [a]

[a] Biological & Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK
[b] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[c] Department of Geography, Durham University, Durham DH1 3LE, UK

HIGHLIGHTS

- The ViPER model is applied to two contrasting catchments.
- Outputs represent spatially distributed maps of predicted *E. coli* burden to land.
- Model shows how *E. coli* burden varies by catchment type and land use composition.
- ViPER enables spatially targeted decision-making for managing *E. coli* in catchments.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 27 September 2017
Received in revised form 22 October 2017
Accepted 24 October 2017
Available online 27 October 2017

Editor: D. Barcelo

Keywords:
Agriculture
Decision-making
Diffuse pollution
Faecal indicator organism
Risk screening

ABSTRACT

Effective management of diffuse microbial water pollution from agriculture requires a fundamental understanding of how spatial patterns of microbial pollutants, e.g. *E. coli*, vary over time at the landscape scale. The aim of this study was to apply the **Vi**sualising **P**athogen & **E**nvironmental **R**isk (ViPER) model, developed to predict *E. coli* burden on agricultural land, in a spatially distributed manner to two contrasting catchments in order to map and understand changes in *E. coli* burden contributed to land from grazing livestock. The model was applied to the River Ayr and Lunan Water catchments, with significant correlations observed between area of improved grassland and the maximum total *E. coli* per 1 km^2 grid cell (Ayr: $r = 0.57$; $p < 0.001$, Lunan: $r = 0.32$; $p < 0.001$). There was a significant difference in the predicted maximum *E. coli* burden between seasons in both catchments, with summer and autumn predicted to accrue higher *E. coli* contributions relative to spring and winter ($P < 0.001$), driven largely by livestock presence. The ViPER model thus describes, at the landscape scale, spatial nuances in the vulnerability of *E. coli* loading to land as driven by stocking density and livestock grazing regimes. Resulting risk maps therefore provide the underpinning evidence to inform spatially-targeted decision-making with respect to managing sources of *E. coli* in agricultural environments.

© 2017 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Diffuse water pollution from agriculture represents a significant threat to the water quality and biodiversity of aquatic ecosystems

* Corresponding author at: Biological & Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Scotland, UK.
E-mail address: david.oliver@stir.ac.uk (D.M. Oliver).

https://doi.org/10.1016/j.scitotenv.2017.10.263
0048-9697/© 2017 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

around the world (Flávio et al., 2017). Spatially targeted decision-making and deployment of mitigation is therefore critical for effective and efficient water resource management, thus helping to reduce agricultural impacts on surface waters (Vinten et al., 2017; Greene et al., 2015). Developments in nutrient management planning and efforts to limit nitrogen (N) and phosphorus (P) export from land to water have highlighted the importance of critical source areas (CSAs), defined as zones in the landscape where high sources of nutrients coincide with high potential for hydrological transfer (Heathwaite et al., 2000). The CSA approach thus represents a powerful tool for conceptualising pollutant risk in agricultural systems and, importantly, it is not restricted to the management of nutrient pollution. Indeed, a significant proportion of surface water contamination with faecal indicator organisms (FIOs), of which *Escherichia coli* is one of the most common, can be attributed to CSAs of microbial pollution (Oliver et al., 2012).

The detection of waterborne *E. coli*, and other FIOs (such as intestinal enterococci), indicates that at some point faecal material has been delivered to the aquatic environment. This delivery can occur via diffuse inputs from agricultural or urban runoff, from point source inputs such as leaking septic tanks or sewage outflows, from the direct deposition of faeces into receiving waters from livestock, or from wildlife that commonly frequent riparian corridors, e.g. deer, birds (Neill et al., 2018; Pattis et al., 2017; Schijven et al., 2015). Much effort has focused on targeting point source inputs, given that this 'end-of-pipe' spatial location is often easily identifiable. In contrast, the management of diffuse and wildlife contributions of microbial pollution to receiving waters represents a more difficult challenge for the policy and regulatory community.

There are examples of national-scale screening tools and models that attempt to identify relative contributions of microbial loads from catchments under high and low river discharge, and the importance of urban versus agricultural land use (Kay et al., 2010; Palazón et al., 2017). These studies provide useful information at the national level as to where and why FIO loads are likely to be generated, and help to highlight potential areas of coastline that are vulnerable to microbial pollution from contaminated catchment discharge. However, such screening tools do not provide information on the nuances in spatial variation of where *within* a catchment FIOs are likely to originate (Dymond et al., 2016). They cannot, therefore, guide finer-scale decision-making concerning where management and mitigation should be prioritised to reduce delivery of FIOs from land to water. Clearly, this represents a different spatial scale of interest; approaches that aim to identify high-risk catchments at a national-scale, coupled with the mapping of spatial variations of relative risk across a landscape within those catchments can provide a more integrated, complementary risk assessment (Heathwaite et al., 2005).

The first step in pinpointing potential CSAs of diffuse microbial water pollution associated with grazing ruminant livestock is the spatial identification of high FIO sources (e.g. *E. coli*) on pasture, hereafter termed burden. The overall burden of *E. coli* derived from grazing activity is governed by the number and type of livestock, and their associated faecal excretion rates and *E. coli* shedding potential (Coffey et al., 2016). In addition, the grazing duration will dictate the rate of replenishment of fresh faeces and associated *E. coli* content to pasture. Post defecation, the concentration of *E. coli* is influenced by a number of environmental factors (Tian et al., 2002), e.g. temperature, UV irradiance and rainfall, which can directly moderate the moisture content of the faecal habitat and indirectly impact on availability of nutrients within the faecal matrix (Oliver and Page, 2016). Ultimately, the coupling of spatial patterns of *E. coli* burden hotspots with an understanding of the likelihood of cell mobilisation from faeces following rainfall and their onward hydrological transfer through the catchment is an essential pre-requisite to enable identification of grazed pasture most vulnerable for contributing diffuse microbial pollution to water (Oliver et al., 2010).

The development of a catchment-scale model of *E. coli* burden is therefore required to map spatial patterns of *E. coli* accumulation on land, but can be challenging given the potential variability in rates of faecal excretion and *E. coli* shedding across and within livestock types. Others have avoided this issue by attributing a set value of *E. coli* shedding per livestock unit, with no differentiation of *E. coli* content or die-off associated with different livestock types (Tian et al., 2002). However, logical structuring of a simple empirical model accounting for differential *E. coli* behaviour by livestock type, which can be applied spatially, will provide a first approximation of relative risk of *E. coli* accumulation across agricultural catchments. In response, the overall aim of this study was to evaluate the recently developed **Vi**sualising **P**athogen & **E**nvironmental **R**isk (ViPER) model, which was created to predict *E. coli* burden on agricultural land (Oliver et al., 2017). The objectives of this study were to: (i) apply the ViPER model in a spatially distributed manner to two contrasting catchments where livestock numbers were known to represent relatively high and low grazing densities; and (ii) use outputs from the model to explore where, when and why *E. coli* burden varies within these two contrasting catchments, thus demonstrating value to catchment managers.

2. Materials and methods

2.1. Case study catchments

To demonstrate the potential of the ViPER model to operate in a spatially distributed manner, we selected two contrasting test catchments in Scotland (Fig. 1) where knowledge and understanding of agricultural practices are relatively well reported (Aitken, 2003; Vinten et al., 2017). The River Ayr and the Lunan Water catchments were selected because they represent different and distinct agricultural regions, and both have designated EU bathing waters at the catchment outlet. The spatial

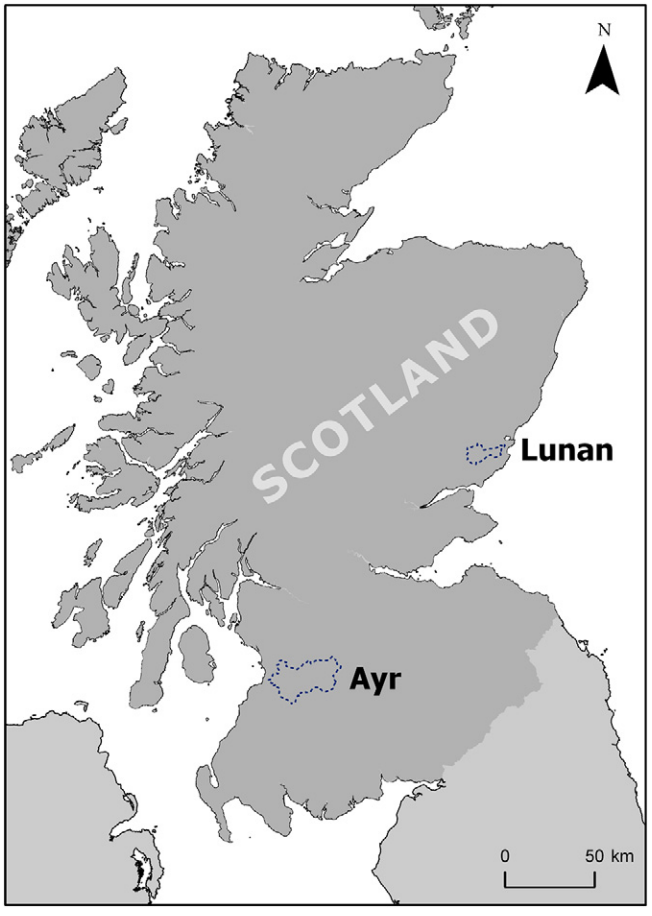


Fig. 1. Spatial locations of the River Ayr catchment and the Lunan Water catchment in Scotland.

pattern and percent coverage of land cover composition is presented in Fig. 2 (Rowland et al., 2017).

2.1.1. The River Ayr catchment

The River Ayr in southwest Scotland drains a catchment area of 574 km^2. Land use in the catchment is predominantly improved grassland, and this area represents a significant proportion of dairy farming in Scotland. Approximately 70% of land area in the River Ayr catchment is used for grazing. The characteristic soils of central and northern Ayrshire are clays and non-calcareous clays with brown forest soils, the maximum elevation is 590 m and the annual average rainfall is 1200 mm. The catchment discharges into an area of coastline that accommodates a series of designated EU bathing waters, which often experience poor microbiological water quality (SEPA, 2017).

2.1.2. The Lunan Water catchment

The Lunan Water catchment is located in the East of Scotland and drains an area of 124 km^2. Land in the catchment is predominantly used for arable and cereal crops, though it represents a typical mixed arable farmland catchment with 15% of land area used for grazing. The underlying soil is of a sandy texture, the maximum elevation in the catchment is 250 m and the annual average rainfall is ~820 mm. The catchment drains into the Lunan Bay, a designated bathing water, which regularly achieves a 'good' classification for microbial water quality, although in 2014 it failed to achieve even the 'sufficient standard' for the first time in its monitored history (SEPA, 2017).

2.2. The ViPER model

ViPER has evolved as an empirical model first reported as part of a cross-disciplinary toolkit for assessing farm-scale contributions to *E. coli* risk (Oliver et al., 2009), which has since been developed and refined (Oliver et al., 2010, 2012). Briefly, ViPER is constructed by using biological parameters of faecal excretion and the *E. coli* shedding and die-off rate. Parameter values for daily *E. coli* shedding by dairy cows, beef cows, calves, sheep and lambs are included in the model (Table 1) and can be set to represent local conditions where data are available. The model accounts dynamically for the accumulation and depletion of *E. coli* burden to land at daily time-steps. The concentration of *E. coli* on pasture is calculated as the sum of two terms, (i) the daily fresh input of *E. coli* by livestock defecation; and (ii) the *E. coli* burden from previous days, which is declining as a result of first-order die-off:

$$E_{(x)} = Ein_{(x)} + E_{(x-1)}e^{-bx} \quad (1)$$

where E_x is the magnitude of the *E. coli* store on day x, Ein is the *E. coli* input of fresh deposits, *e* is a mathematical constant (base of natural log), b is the exponential die-off constant. Eq. (1) is applicable to each livestock type present, with the sum of all livestock types representing the total *E. coli* burden over a specified number of days, i.e. the grazing period. Specifically, daily *E. coli* loading is calculated by multiplying the number of livestock by both the daily dry matter excreted per livestock type and a typical value for *E. coli* per gram of dry faeces associated with each livestock type (Table 1).

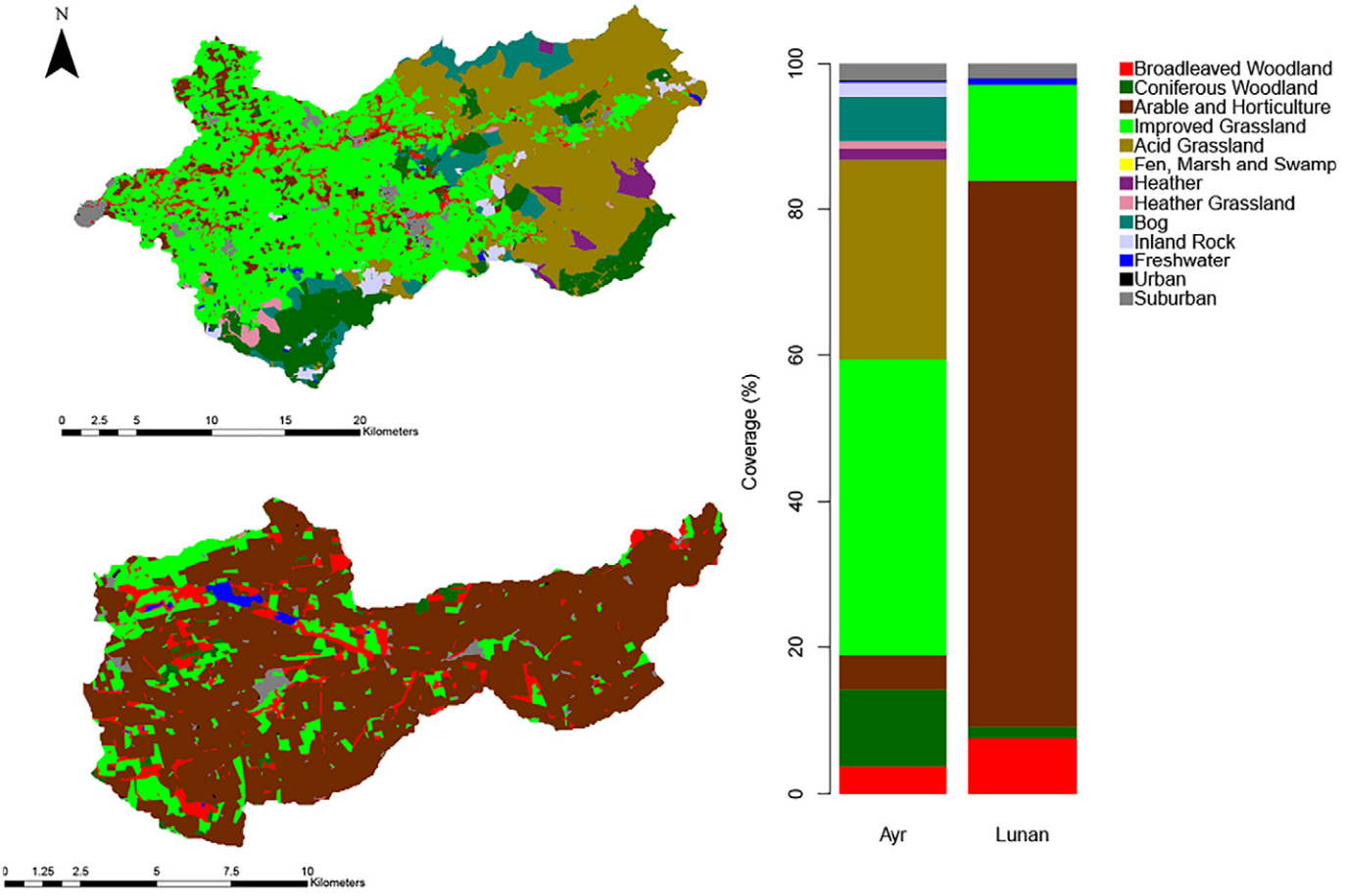


Fig. 2. Land cover in the River Ayr (upper map) and Lunan Water (lower map) catchments – stacked bars show the percentage land cover composition of both catchments. Features of this map are based on digital spatial data (Rowland et al., 2017).

Table 1
Parameters used for different livestock types in the ViPER model applied to the River Ayr and Lunan Water catchments.

Livestock type	Daily *E. coli* burden (CFU)	Min die-off rate (day^{-1})	Max die-off rate (day^{-1})
Dairy cow	1.8×10^{10}[a]	−0.0606[e]	−0.0909[c]
Beef cow	4.2×10^{10}[b]	−0.0606[e]	−0.0909[c]
Calf	2.1×10^{10}[c]	−0.0606[e]	−0.0909[c]
Sheep	5.0×10^{9}[d]	−0.0640[e]	−0.0920[c]
Lamb	1.01×10^{10}[c]	−0.0640[e]	−0.0920[c]

[a] Oliver and Page (2016).
[b] Oliver et al. (2010).
[c] Oliver et al. (2009).
[d] Hodgson et al. (2009).
[e] Avery et al. (2004).

2.3. Parameterisation and running the ViPER model

Validation of the underlying empirical model has previously been reported to demonstrate that the approach is fit-for-purpose (Oliver et al., 2012). Livestock numbers are a fundamental input parameter for ViPER, allowing the algorithm described by Eq. (1) to calculate *E. coli* burden from fresh deposits. The spatial extent to which these livestock numbers have access to grazing is also important to enable an *E. coli* burden per unit area to be determined. Over 8000 land parcels (i.e. field boundaries) and their associated livestock numbers linked to farm businesses within each catchment were obtained from the Scottish Government. Data were available via entries from all farm businesses in 2015 that had successfully completed and submitted a Single Application Form (SAF) to Rural Payments & Services which enables the integrated administration and control system (IACS) to produce an annual register as part of the 'Basic Payment Scheme' (the largest of the EU's rural grants and payments). IACS data enabled livestock to be categorised into dairy and beef cows, calves, sheep and lambs. Working with spatial data collected via survey returns such as those described above can present a number of challenges. For example, IACS only provides details of the livestock (number and type) that a producer has present on 1st March of each year; however, these figures represent producer rather than holding level, and potentially a producer may have multiple holdings, some of which fall outside of the specific catchment boundaries of interest. IACS does not provide information on which location/holding the livestock are physically associated with, and although such cases are likely to be uncommon, it nevertheless represents a limitation of spatial data returns of this nature. Furthermore, specific dates on the timing of grazing regimes across different businesses are not available via the IACS dataset. Thus, the following assumptions were made for the purposes of our study: (i) all livestock data were assumed to link to holdings located within the catchment boundaries; (ii) the livestock attributed to each farm were evenly distributed across the combined land area associated with their respective holding; and (iii) typical grazing periods were assumed to be common to all farms, with grazing regimes for bovine and ovine livestock shown in Table 2.

The livestock data obtained via IACS were at the 'business unit' level rather than 'per farm field', therefore the field dataset was merged based on the Business Reference Number (BRN) to create a multi-polygon geometry. The ViPER model was used to determine *E. coli* burden across the Ayr and Lunan catchments at a 1 km^2 resolution. Briefly, field definitions of each farm were combined, even if distant, to form a more complex single geometry record in the database table that could define all of the contributing field boundaries in each catchment (Fig. 3). The model was run over an annual period to calculate the *E. coli* burden ha^{-1} for each calendar day over an annual period for each multi-polygon record (corresponding to each BRN). To preserve farm anonymity, an additional GIS layer was created to obfuscate the results in each catchment. This consisted of creating a fishnet grid layer with cells of 1 km by 1 km, i.e. for the Ayr catchment, 31 by 43 cells were required from the lower left coordinate (235,647,608,627) in British National Grid and for the Lunan catchment, 15 by 23 cells were required from British National Grid coordinate (344,819,742,403). Each grid cell was assigned a unique ID, and the cell geometry used to cookie cut through the farm definitions (1 farm per BRN). The new size of each farm section was calculated in hectares, reflecting any changes from the subdivision of the geometries based on the fishnet grid layer. The total *E. coli* burden within each grid cell was converted to reflect the *E. coli* burden ha^{-1} of farmland (not per 1 km^2) as some cells accommodate land use other than that used for agriculture. Thus, for any day in the calendar we could determine the sum of *E. coli* burden across each grid cell (1 km by 1 km), which may have had *E. coli* contributions from a complete farm business, or parts of a farm business, or even part of a field from another business.

From this dataset, it was possible to calculate the maximum value of *E. coli* burden linked to grazing activity for any given cell over the course of a year (each maximum *E. coli* burden value per cell may have occurred on a different day of the year) and to determine patterns of *E. coli* burden over time. Maps of total burden were captured on the 1st March, 1st June, 1st September and 1st December to demonstrate temporal and spatial variability in *E. coli* and to map onto key differences in land management over the year.

2.4. Statistical analysis

Spatially distributed values of maximum *E. coli* burden associated with both catchments did not meet the assumption of normal distribution after log transformation. Thus, non-parametric statistical tests were used to evaluate the catchment datasets. A Kruskal-Wallis test was used to test for differences between the magnitudes of *E. coli* burden on agricultural land at defined time points through the calendar year (i.e. seasonal differentiation in burden). The Wilcoxon signed-rank test was used to determine whether the total *E. coli* burden ha^{-1} of farmland over a complete annual cycle was significantly different between the two catchments. Spearman's correlation (r) was used to compare the area of improved grassland and the total maximum *E. coli* burden within each 1 km^2 grid cell.

3. Results

The modelled distributions of *E. coli* burden across the River Ayr and Lunan Water catchments are shown in Figs. 4 and 5, respectively, and represent some of the first catchment wide spatially distributed maps of predicted *E. coli* burden to land contributed from grazing livestock. The maximum value of predicted total *E. coli* burden ha^{-1} of farmland for each 1 km^2 grid cell is shown for both catchments (Fig. 4a, Fig. 5a). It should be noted that the maximum occurrence of the total *E. coli* burden within each grid cell is likely to vary temporally depending on the composition of livestock grazing activity. Thus, different grid cells may represent different calendar days in Figs. 4a and 5a. For example, the highest recorded grid cell value for the total predicted *E. coli* burden in the River Ayr catchment registered 13.5 $\log_{10}$ CFU ha^{-1} farmland on day 203 of the year (July 22nd), whereas in the Lunan Water catchment a peak burden of 13.4 $\log_{10}$ CFU ha^{-1} occurred on day 204 of the year (July 23rd). On average, peak *E. coli* burden ha^{-1} farmland occurred on day 203 in both catchments.

Table 2
Assumed grazing periods applied to all livestock in both catchments.

Livestock type	Grazing period
Dairy cow	1st April–31st October
Beef cow	1st April–31st October
Calf	1st April–31st October
Sheep	1st Jan–14th April (removed for lambing); 1st May–31 Dec
Lamb	1st May–1st November

1. Calculate the *E. coli* burden for each Business (e.g. 1,2,3,4) – based on the known livestock numbers. There may be multiple fields per Business. This results as *E. coli* burden per hectare for each Business for each calendar day.
2. Create 1km by 1km fishnet grid, give each grid cell a unique ID
3. Calculate new field geometries by cutting up each business according to fishnet grid squares
4. Calculate area (in Ha) for each Business per grid square
5. Multiply the area of each Business in each grid cell by the calculated *E. coli* / Ha value (step 1) – to get the total amount of *E. coli*
6. Calculate the SUM of *E. coli* burden per hectare per square for each calendar day, and divide by the total farmland in each cell to give the burden per hectare of farmland per grid cell. These are used for the Spring, Summer, Autumn, Winter maps.
7. The MAXIMUM value for each cell throughout the year can be calculated, to reveal the highest burdens spatially.

Note:
Each Business may contribute to many grid squares
Each grid square may have many source inputs
Parts of grid squares may be blank (not farms)

Fig. 3. Steps used to rasterise the ViPER outputs to 1 km^2 grid cells.

The total *E. coli* burden changes spatially through time over the course of a year, with a snapshot of predicted total burden for every grid cell shown for 1st March, 1st June, 1st September and 1st December to represent spring, summer, autumn and winter seasonal burden, respectively (panels b, c, d and e of Figs. 4 and 5). For both catchments, there was a significant difference in the predicted *E. coli* burden between seasons, with modelled data for 1st June and 1st September predicting a higher *E. coli* contribution relative to 1st March and 1st December ($p < 0.001$). In the Lunan and Ayr catchments, the area of farmland accommodating an *E. coli* burden in excess of 3×10^{12} CFU ha^{-1} decreased by over 89% and 80% of the total catchment area, respectively, in the transition from autumn to winter.

The spatial distribution of *E. coli* burden closely reflected the spatial patterns of land use in both catchments (Figs. 2, 4 & 5). The correlation between area of improved grassland and the maximum total *E. coli* per 1 km^2 grid cell was stronger in the River Ayr catchment ($r = 0.57$; $p < 0.001$) compared with the Lunan Water ($r = 0.32$; $p < 0.001$). In total, 4 km^2 of the River Ayr catchment accommodated a total *E. coli* burden in excess of 13.2 $\log_{10}$ *E. coli* ha^{-1} of farmland, all relatively close to the catchment outlet that drains to the west coast. In the Lunan Water catchment 1 km^2 reached an *E. coli* burden of 13.2 $\log_{10}$ *E. coli* ha^{-1} of farmland. In the Lunan Water catchment, the model identified that the largest sources of *E. coli* would accumulate in areas farthest from the catchment outlet, which drains to the east coast. In both catchments, the model did predict sporadic grid cells, representing small areas of land, accommodating high *E. coli* burden for the spring and winter snapshot sample.

The time-series profile of *E. coli* burden contributed from grazing livestock, normalised to a per ha basis over the course of an annual period, is represented for both catchments in Fig. 6. This profile was derived using all IACS data returned for livestock numbers combined with the assumed grazing durations used in the ViPER model (Table 2). This result represents a first approximation of how *E. coli* burden would change over time in both catchments, as driven by direct defecation from livestock, prior to any hydrological mobilisation and transfer of cells.

Based on the spatial data associated with the IACS returns for each catchment, the total area of farmed land that was modelled within the River Ayr catchment was 40,250 ha, compared with 6799 ha for the Lunan Water catchment. Accepting the model caveat that predictions are based on total *E. coli* burden contributed to land independent of any loss from hydrological processes such as runoff following rainfall events, the total burden across all farmed land for a given year, based on livestock stocking information from IACS returns was 15.6 $\log_{10}$ *E. coli* and 16.6 $\log_{10}$ *E. coli*, for the Lunan and Ayr catchments, respectively. The order of magnitude difference in burden corresponds to the much higher overall proportion of improved grassland in the River Ayr catchment (Fig. 2). Both catchments reached the maximum for their respective total *E. coli* burden on day 203 of the year (July 22nd) based on the assumed grazing regimes. When normalised for the total amount of farmed land within each catchment the maximum *E. coli* burden was equivalent to, on average, 12.0 $\log_{10}$ *E. coli* ha^{-1} and 11.8 $\log_{10}$ *E. coli* ha^{-1}, for the Ayr and Lunan catchments, respectively. Over the course of an entire annual period the normalised *E. coli* burden modelled in the River Ayr catchment was significantly higher than that modelled in the Lunan Water catchment ($p < 0.001$).

4. Discussion

This study reports on the successful deployment of the ViPER model to two contrasting catchments in Scotland, providing spatially distributed predictions of *E. coli* burden through time. Despite the obvious importance to catchment management, little research has specifically reported on the modelling of *E. coli* burden at the landscape scale, nor has mapped the subtle complexity in patterns of *E. coli* burden on land that manifest as a result of different livestock grazing regimes across catchments (Oliver et al., 2012). By using two contrasting exemplar catchments, our modelling results show how the distribution of *E. coli* burden contributed from grazing livestock varies by catchment type and land use composition and how overall burden will vary through time.

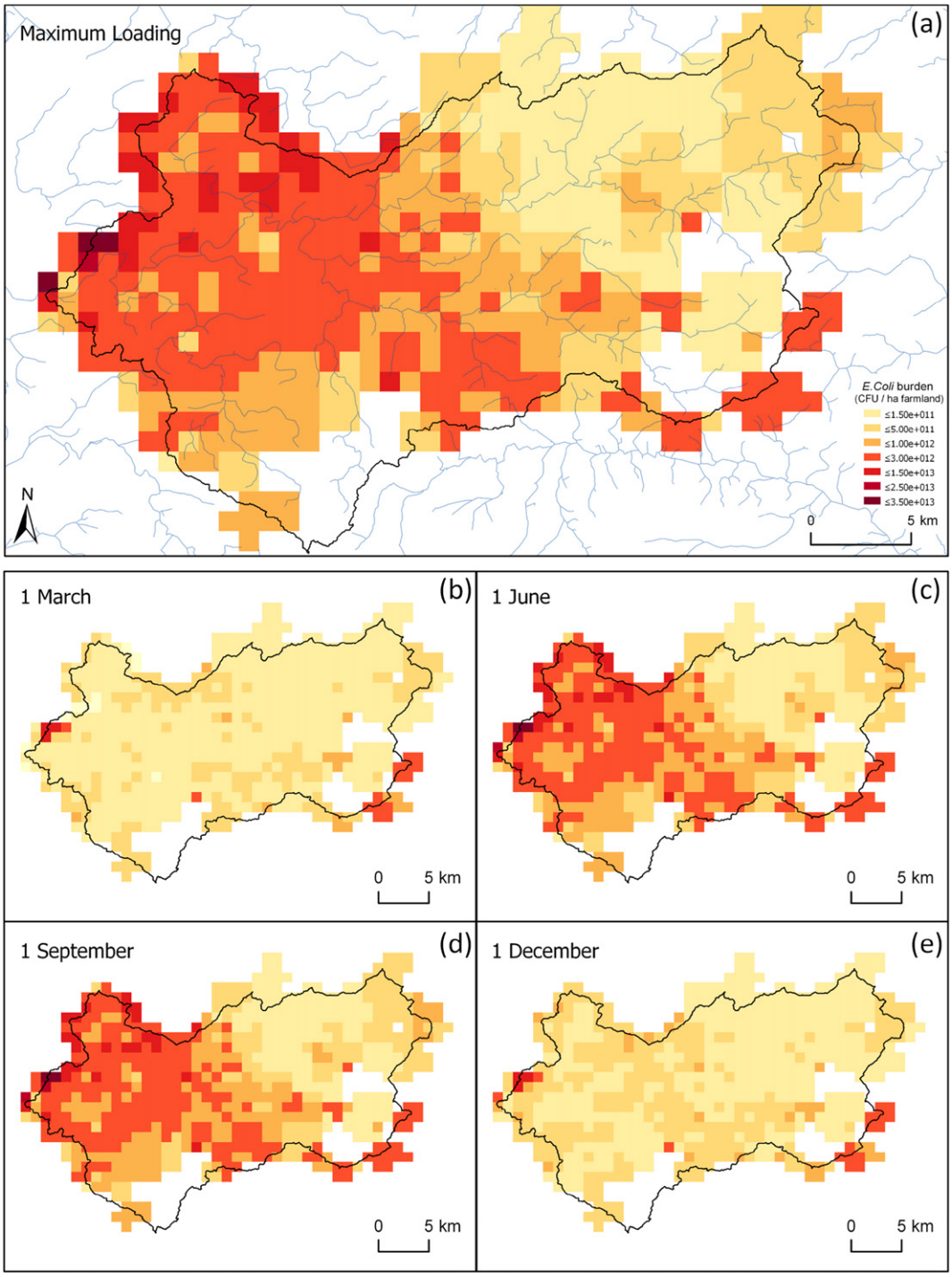


Fig. 4. Modelled patterns of maximum *E. coli* burden ha^{-1} of farmland distributed across the River Ayr catchment at 1 km^2 grid resolution. Panel A represents the peak occurrence of burden within each grid cell; B, C, D & E show maximum burden for specific snapshots in time. White cells represent no data (no farms or no data about livestock on the farms).

The application of a relatively simple underlying empirical model to 1 km^2 grid cells spanning the River Ayr and Lunan Water catchments demonstrates the potential transferability of this approach to the rest of the UK and beyond with similar climate and livestock management. However, some modification of die-off parameters to reflect local temperature and UV regimes across seasons would be beneficial to further enhance model transferability. Thus, ViPER provides a landscape screening tool to identify high-risk zones for the accumulation of *E. coli*, as has been done for other diffuse agricultural pollutants such as P (Strömqvist et al., 2008; Heathwaite et al., 2003). Furthermore, the visualisation of relative differences in *E. coli* burden across a catchment, rather than the absolute values predicted by the model, is likely to be of practical use to policy- and decision-makers because it enables spatial prioritisation of effort with respect to managing sources of FIOs in catchments.

The predictions reported in this study do not account for land applications of FIOs via spreading of organic manures such as solid farmyard manure (FYM) or liquid slurries. However, the relative contribution of *E. coli* from managed manures is known to be orders of magnitude less than from the fresh faeces of grazing livestock (Vinten et al., 2004). Likewise, contributions of *E. coli* from wildlife faeces are not included in ViPER because more evidence is needed to understand better the fate

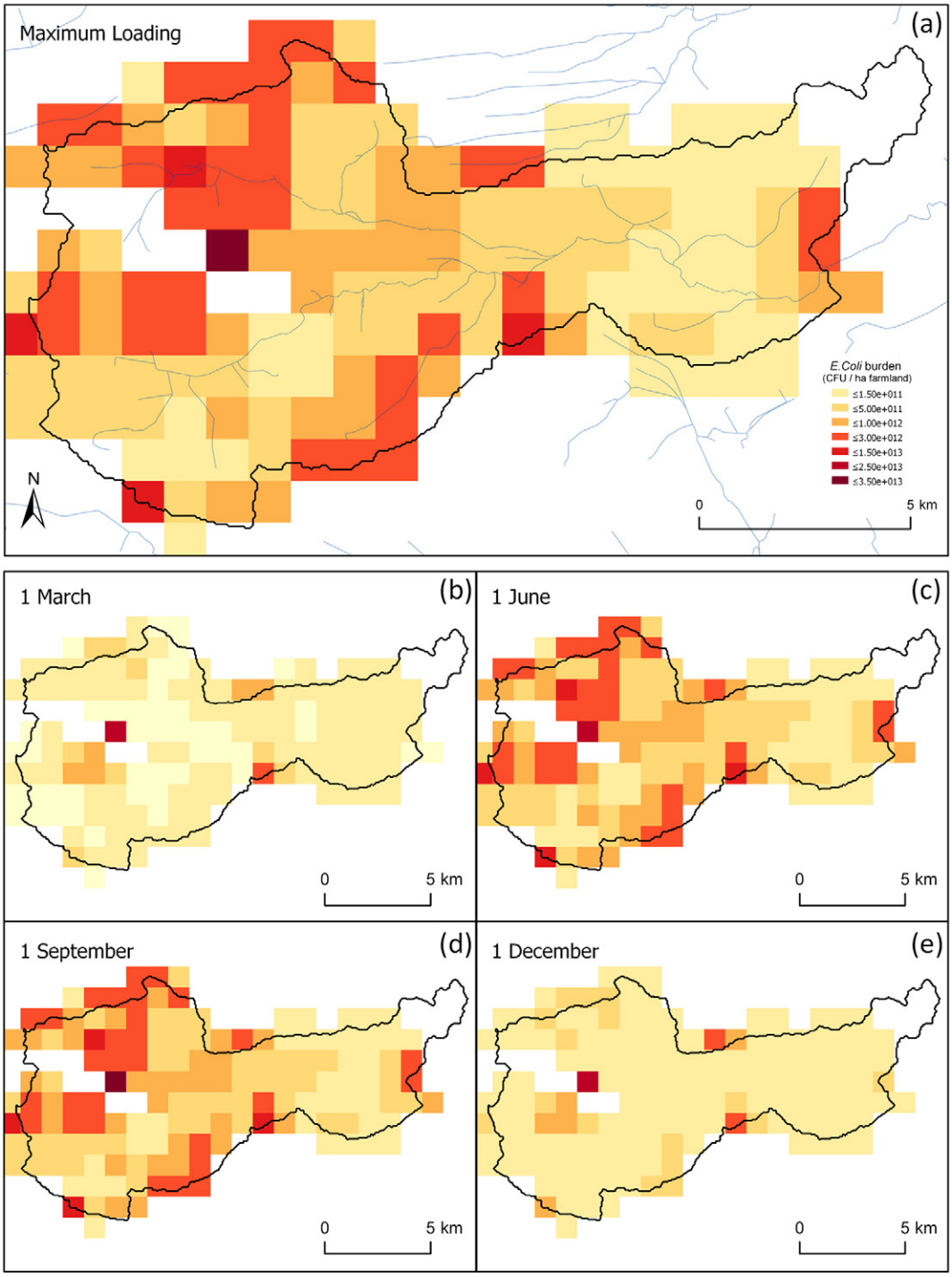


Fig. 5. Modelled patterns of maximum *E. coli* burden ha^{-1} of farmland distributed across the Lunan Water catchment at 1 km^2 grid resolution. Panel A represents the peak occurrence of burden within each grid cell; B, C, D & E show maximum burden for specific snapshots in time. White cells represent no data (no farms or no data about livestock on the farms).

and release of *E. coli* from such sources (Guber et al., 2015). Irrespective of these omissions, catchment-wide predictions of *E. coli* burden from livestock grazing provide a useful mechanism for promoting knowledge exchange with catchment managers and raising awareness as to where in a catchment high source loading of *E. coli* may occur (Oliver et al., 2017). Evaluating how future shifts in livestock management might impact on *E. coli* burden is one potential application of ViPER that would be advantageous to catchment managers. For example, scenarios of extended cattle grazing periods during drier weather or shifts to permanent housing of dairy cattle in wetter areas of the UK can be modelled to demonstrate relative spatial and temporal differences in resulting *E. coli* burden on pasture, and be used as evidence to support farm management decisions (Jones et al., 2017). Impacts of climate change could be considered through varying the die-off coefficients used in the underpinning model. Likewise, if microbial water quality at designated sampling points (e.g. bathing waters at coastal or inland locations) is below regulatory standards then the application of ViPER to the contributing catchment will highlight particularly high burden areas for subsequent investigation by environmental regulators. Given that a core function of this model is to provide a rapid initial screen to understand relative distributions of FIO sources in a catchment, the use of assumed (rather than specific) grazing dates attributed in unison

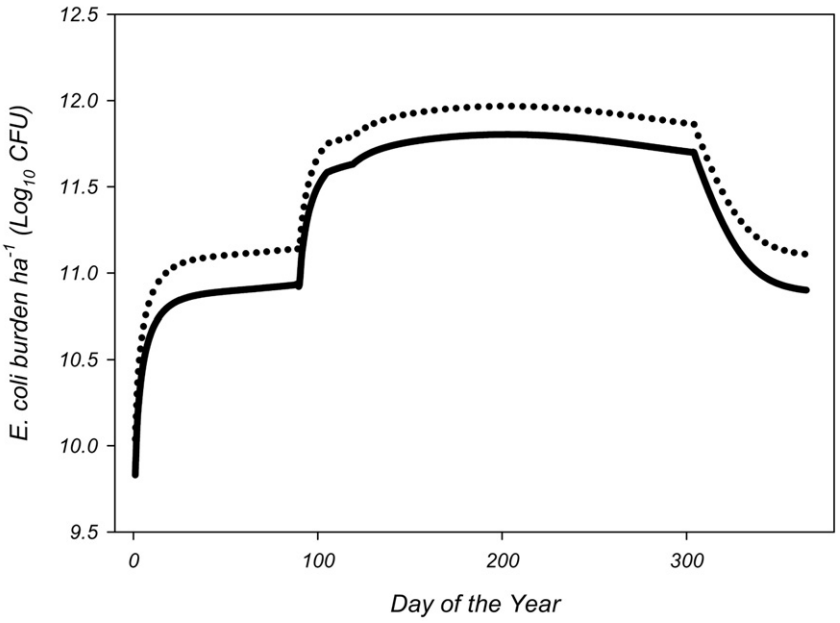


Fig. 6. Modelled time-series of *E. coli* burden per ha of farmed land in the River Ayr catchment (dashed line) and Lunan Water catchment (solid line) over an annual cycle, based on livestock data provided through farm survey returns.

to all farms to generate predictions of *E. coli* burden is considered appropriate. The investment of effort required to deploy targeted custom farm surveys to obtain specific dates for grazing periods in order to fine-tune the model is probably a poor use of resources relative to the likely extent of improvements it would deliver in modelled predictions.

Of those models that predict FIO risk to the water environment, many do include sub-modules that account for *E. coli* stores in the landscape. For example, the Soil and Water Assessment Tool (SWAT) has been modified to predict *E. coli* concentrations in receiving waters and includes a sub-routine to account for FIO loading to land (Coffey et al., 2010). Similarly, the Bacteria Source Load Calculator (BSLC) was developed in the USA and includes an inventory of livestock for each catchment in order to generate estimates of *E. coli* burden, which are then used as input files in models such as the Hydrological Simulation Program Fortran (HSPF) (Zeckoski et al., 2005). Others have developed deterministic models for catchments and incorporated a land budget module to account for microbial loading from livestock to then distribute through a catchment system following mobilisation from rainfall (Ferguson et al., 2005). These microbial input sub-modules are built on empirical data associated with observed microbial die-off rates and loading in faeces. However, an assumption of those models is that input parameters associated with grazing livestock combine to accurately reflect the overall *E. coli* burden contributed to land over time in a given area of study, but this remains untested.

In marked contrast to these models, the underpinning empirical burden model that informs ViPER has been previously tested at the headwater catchment scale (0.4 km^2) and was able to satisfactorily predict *E. coli* burden on pasture, with 89% of observed values falling within the minimum and maximum range of predicted values (Oliver et al., 2012). This result provides a level of confidence that the deployment of the ViPER model in a distributed manner across larger catchment areas is transferable and can return credible predictions of *E. coli* burden. We are not aware of any other *E. coli* model that has had the source loading component verified at the landscape scale to ascertain whether the die-off functions used are fit-for-purpose, and therefore the spatial model we present is associated with a degree of proven robustness that is not afforded to many other models of catchment scale FIO dynamics. Of course like any model, each calculation stage in our modelling process has associated uncertainty, which can be attributed to both the spatial input data and the model calculations and coefficients (Hamilton et al., 2015). However, it is difficult to fully verify the model output versus real world observations, as this would require measurements of the volume of faeces excreted onto land within a large number of the 1 km^2 grid cells and assessments of the associated *E. coli* concentrations in faecal material. To undertake this form of model validation would be extremely labour intensive, as would be the case for cell-by-cell verification of any (semi) distributed model.

Using the two exemplar catchments of the River Ayr and Lunan Water, the areal extent of improved grassland appears to clearly explain patterns of high *E. coli* burden. The majority of land in the River Ayr catchment is improved grassland, which tends to carry a greater density of livestock and consequently these areas also accommodate the greatest *E. coli* burden derived from predominantly dairy, but also beef, faeces (Crowther et al., 2002). The ViPER model does not directly associate grassland areas with livestock, although livestock data (derived from IACS) are distributed equally across all of the fields linked to each farm business. It is likely that many farm businesses that hold livestock will be registered as holdings within those areas of the catchment where grassland is a predominant land cover. Therefore, we apportioned grazing activity by dividing the total head of livestock per farm by the corresponding farm area in the catchment, but in reality certain fields are likely to be much more heavily grazed while others may never be stocked with sheep or cattle (Jakoby et al., 2014). Access to farmer records of livestock movement data and field-by-field stocking numbers at catchment scales is notoriously difficult to obtain and is a common factor leading to the use of model assumptions (Dymond et al., 2016). In addition, such data is likely to vary year-on-year. Apportioning livestock to farm fields could potentially be improved by differentiating between fields whose boundaries cut across grassland versus 'other' land cover types, with livestock then attributed to the former only. The isolated, sporadically high burden values that were occasionally predicted may reflect another limitation of the model. These likely provide an example of where a farm holds a particular livestock type in large quantities (e.g. sheep) but owns large areas of land (e.g. rough grazing) in a location beyond the immediate catchment of interest. The model cannot accurately reflect this at the catchment scale and so apportions all livestock to a potentially small area of land linked to the associated farm within the modelled catchment. Modelling a much larger area such as the whole of Scotland, rather than individual

catchments, would probably eliminate this error, associated with an artefact of scaling (Winter et al., 2011).

The differences in the historical classifications of bathing water quality associated with the designated bathing waters into which these two catchments drain, would appear to complement the predicted magnitude and spatial pattern of *E. coli* burden for the two catchments. For example, the coastal waters of Ayrshire are more frequently impaired by microbial pollution and the ViPER model clearly highlights a larger catchment source of *E. coli* in closer proximity to the catchment outlet. However, the modelled outputs from ViPER predict differences in burden only, and do not differentiate any spatial variability in soil type, rainfall or hydrological connectivity of the River Ayr or Lunan Water catchments. Efforts to link these predictions of *E. coli* burden to hydrological risk-mapping tools, e.g. SCIMAP (Reaney et al., 2011), could provide a powerful coupling to inform on FIO delivery from land to water across agricultural landscapes and would provide the next logical step to develop this approach into a model for predicting FIO risk to the stream network (Bergion et al., 2017). This would allow for both components of CSAs (burden and connection to receiving waters) to be represented. Similar frameworks have evolved to predict the risk of diffuse P loss from land to water, incorporating multiple spatial layers to account not only for indicators of P source but also P transfer and P delivery (Heathwaite et al., 2003); however, in the case of P our level of understanding concerning transfer and delivery processes is more advanced than for FIOs (Kay et al., 2010).

5. Conclusion

The ViPER model describes, at the 1 km grid cell scale, how catchment wide vulnerability to *E. coli* loading on land will vary according to stocking density and livestock grazing regimes. To quantify the risk of *E. coli* burden to receiving waters, the next steps in model development must focus on coupling spatial predictions of *E. coli* burden to an understanding of cell mobilisation and onward hydrological connectivity to the stream network. Doing so would increase the likelihood that such a tool could be deployed by the relevant regulatory and environmental agencies in order to spatially target management and remediation efforts, and ultimately protect bathing water environments vulnerable to microbial pollution. Therefore, ViPER represents a key development in what must be a continued effort to improve spatial and temporal risk mapping of *E. coli* risks within agricultural catchments.

Acknowledgements

This research was supported by the UK Natural Environment Research Council, as part of the **Vi**sualising **P**athogen & **E**nvironmental **R**isk (ViPER) projects (NE/M007812/1 and NE/P016332/1). The work contains Scottish Government information licensed under the Open Government License v3.0. We thank the three reviewers for their constructive comments on our manuscript.

References

Aitken, M.N., 2003. Impact of agricultural practices and river catchment characteristics on river and bathing water quality. Water Sci. Technol. 48, 217–224.
Avery, S.M., Moore, A., Hutchison, M.L., 2004. Fate of *Escherichia coli* originating from livestock faeces deposited directly onto pasture. Lett. Appl. Microbiol. 38, 355–359.
Bergion, V., Sokolova, E., Åström, J., Lindhe, A., Sörén, K., Rosén, L., 2017. Hydrological modelling in a drinking water catchment area as a means of evaluating pathogen risk reduction. J. Hydrol. 544, 74–85.
Coffey, R., Cummins, E., O'Flaherty, V., Cormican, M., 2010. Pathogen sources estimation and scenario analysis using the Soil and Water Assessment Tool (SWAT). Hum. Ecol. Risk. Assess. 16, 913–933.
Coffey, R., Benham, B., Wolfe, M.L., Dorai-Raj, S., Bhreathnach, N., O'Flaherty, V., Cormican, M., Cummins, E., 2016. Sensitivity of streamflow and microbial water quality to future climate and land use change in the West of Ireland. Reg. Environ. Chang. 16, 2111–2128.
Crowther, J., Kay, D., Wyer, M.D., 2002. Faecal-indicator concentrations in waters draining lowland pastoral catchments in the UK: relationships with land use and farming practices. Water Res. 36, 1725–1734.
Dymond, J.R., Serezat, D., Ausseil, A.G.E., Muirhead, R.W., 2016. Mapping of *Escherichia coli* sources connected to waterways in the Ruamahanga catchment, New Zealand. Environ. Sci. Technol. 50, 1897–1905.
Ferguson, C.M., Croke, B., Ashbolt, N.J., Deere, D.A., 2005. A deterministic model to quantify pathogen loads in drinking water catchments: pathogen budget for the Wingecarribee. Water Sci. Technol. 52, 191–197.
Flávio, H.M., Ferreira, P., Formigo, N., Svendsen, J.C., 2017. Reconciling agriculture and stream restoration in Europe: a review relating to the EU Water Framework Directive. Sci. Total Environ. 596, 378–395.
Greene, S., Johnes, P.J., Bloomfield, J.P., Reaney, S.M., Lawley, R., Elkhatib, Y., Freer, J., Odoni, N., Macleod, C.J., Percy, B., 2015. A geospatial framework to support integrated biogeochemical modelling in the United Kingdom. Environ. Model Softw. 68, 219–232.
Guber, A.K., Fry, J., Ives, R.L., Rose, J.B., 2015. *Escherichia coli* survival in, and release from, white-tailed deer feces. Appl. Environ. Microbiol. 81, 1168–1176.
Hamilton, S.H., ElSawah, S., Guillaume, J.H., Jakeman, A.J., Pierce, S.A., 2015. Integrated assessment and modelling: overview and synthesis of salient dimensions. Environ. Model Softw. 64, 215–229.
Heathwaite, L., Sharpley, A., Gburek, W., 2000. A conceptual approach for integrating phosphorus and nitrogen management at watershed scales. J. Environ. Qual. 29, 158–166.
Heathwaite, A.L., Fraser, A.I., Johnes, P.J., Hutchins, M., Lord, E., Butterfield, D., 2003. The phosphorus indicators tool: a simple model of diffuse P loss from agricultural land to water. Soil Use Manag. 19, 1–11.
Heathwaite, A.L., Dils, R.M., Liu, S., Carvalho, L., Brazier, R.E., Pope, L., Hughes, M., Phillips, G., May, L., 2005. A tiered risk-based approach for predicting diffuse and point source phosphorus losses in agricultural areas. Sci. Total Environ. 344, 225–239.
Hodgson, C.J., Bulmer, N., Chadwick, D.R., Oliver, D.M., Heathwaite, A.L., Fish, R.D., Winter, M., 2009. Establishing relative release kinetics of faecal indicator organisms from different faecal matrices. Lett. Appl. Microbiol. 49, 124–130.
Jakoby, O., Quaas, M.F., Müller, B., Baumgärtner, S., Frank, K., 2014. How do individual farmers' objectives influence the evaluation of rangeland management strategies under a variable climate? J. Appl. Ecol. 51, 483–493.
Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., 2017. Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. Agric. Syst. 269–288.
Kay, D., Anthony, S., Crowther, J., Chambers, B.J., Nicholson, F.A., Chadwick, D., Stapleton, C.M., Wyer, M.D., 2010. Microbial water pollution: a screening tool for initial catchment-scale assessment and source apportionment. Sci. Total Environ. 408, 5649–5656.
Neill, A.J., Tetzlaff, D., Strachan, N.J.C., Hough, R.L., Avery, L.M., Watson, H., Soulsby, C., 2018. Using spatial-stream-network models and long-term data to understand and predict dynamics of faecal contamination in a mixed land-use catchment. Sci. Total Environ. 612, 840–852.
Oliver, D.M., Page, T., 2016. Effects of seasonal meteorological variables on *E. coli* persistence in livestock faeces and implications for environmental and human health. Sci Rep 6.
Oliver, D.M., Fish, R.D., Hodgson, C.J., Heathwaite, A.L., Chadwick, D.R., Winter, M., 2009. A cross-disciplinary toolkit to assess the risk of faecal indicator loss from grassland farm systems to surface waters. Agric. Ecosyst. Environ. 129 (4), 401–412.
Oliver, D.M., Page, T., Heathwaite, A.L., Haygarth, P.M., 2010. Re-shaping models of *E. coli* population dynamics in livestock faeces: increased bacterial risk to humans? Environ. Int. 36, 1–7.
Oliver, D.M., Page, T., Zhang, T., Heathwaite, A.L., Beven, K., Carter, H., McShane, G., O' Keenan, P., Haygarth, P.M., 2012. Determining *E. coli* burden on pasture in a headwater catchment: combined field and modelling approach. Environ. Int. 43, 6–12.
Oliver, D.M., Bartie, P.J., Heathwaite, A.L., Pschetz, L., Quilliam, R.S., 2017. Design of a decision support tool for visualising *E. coli* risks on agricultural land using a stakeholder-driven approach. Land Use Policy 66, 227–234.
Palazón, A., López, I., Aragonés, L., Villacampa, Y., Navarro-González, F.J., 2017. Modelling of *Escherichia coli* concentrations in bathing water at microtidal coasts. Sci. Total Environ. 593, 173–181.
Pattis, I., Moriarty, E., Billington, C., Gilpin, B., Hodson, R., Ward, N., 2017. Concentrations of spp., Enterococci, and spp. in the feces of farmed red deer in New Zealand. J. Environ. Qual. 46, 819–827.
Reaney, S.M., Lane, S.N., Heathwaite, A.L., Dugdale, L.J., 2011. Risk-based modelling of diffuse land use impacts from rural landscapes upon salmonid fry abundance. Ecol. Model. 222, 1016–1029.
Rowland, C.S., Morton, R.D., Carrasco, L., McShane, G., O'Neil, A.W., Wood, C.M., 2017. Land Cover Map 2015 (Vector, GB). NERC Environmental Information Data Centre https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73.
Schijven, J., Derx, J., de Roda Husman, A.M., Blaschke, A.P., Farnleitner, A.H., 2015. QMRAcatch: microbial quality simulation of water resources including infection risk assessment. J. Environ. Qual. 44, 1491–1502.
SEPA, 2017. Bathing Waters. Available at https://apps.sepa.org.uk/bathingwaters/. (Accessed 17th October 2017).
Strömqvist, J., Collins, A.L., Davison, P.S., Lord, E.I., 2008. PSYCHIC–a process-based model of phosphorus and sediment transfers within agricultural catchments. Part 2. A preliminary evaluation. J. Hydrol. 350, 303–316.
Tian, Y.Q., Gong, P., Radke, J.D., Scarborough, J., 2002. Spatial and temporal modeling of microbial contaminants on grazing farmlands. J. Environ. Qual. 31, 860–869.

Vinten, A.J.A., Douglas, J.T., Lewis, D.R., Aitken, M.N., Fenlon, D.R., 2004. Relative risk of surface water pollution by *E. coli* derived from faeces of grazing animals compared to slurry application. Soil Use Manag. 20, 13–22.

Vinten, A., Sample, J., Ibiyemi, A., Abdul-Salam, Y., Stutter, M., 2017. A tool for cost-effectiveness analysis of field scale sediment-bound phosphorus mitigation measures and application to analysis of spatial and temporal targeting in the Lunan Water catchment, Scotland. Sci. Total Environ. 586, 631–641.

Winter, M., Oliver, D.M., Fish, R., Heathwaite, A.L., Chadwick, D., Hodgson, C., 2011. Catchments, subcatchments and private spaces: scale and process in managing microbial pollution from source to sea. Environ. Sci. Pol. 14, 315–326.

Zeckoski, R.W., Benham, B.L., Shah, S.B., Wolfe, M.L., Brannan, K.M., Al-Smadi, M., Dillaha, T.A., Mostaghimi, S., Heatwole, C.D., 2005. BSLC: a tool for bacteria source characterization for watershed management. Appl. Eng. Agric. 21, 879–889.